心理学报 2021, Vol. 53, No. 1, 55−66
Acta Psychologica Sinica

https://dx.doi.org/10.3724/SP.J.1041.2021.00055

右侧颞顶联合区及道德加工脑网络的功能连接预测社会性框架效应：来自静息态功能磁共振的证据

崔 芳[1,2] 杨佳苗[1] 古若雷[3,4] 刘 洁[1,2]

([1]深圳大学心理学院; [2]深圳大学脑疾病与认知科学研究中心, 深圳 518060)
([3]中国科学院心理研究所行为科学重点实验室, 北京 100101)
([4]中国科学院大学心理学系, 北京 100049)

摘 要 社会性框架效应指个体进行社会性决策时, 描述方式(即框架)对其选择倾向的影响。采用“伤害/帮助框架”范式, 本研究在行为上发现了显著的框架效应：被试在伤害框架下比帮助框架下, 更倾向于选择利他选项, 这可能是由于伤害框架凸显了伤害他人的后果及有意性, 从而提高了道德冲突水平。在神经活动上发现(1)静息态下, 右侧颞顶联合区(TPJ)的低频振荡振幅(ALFF)与框架效应强度存在显著正相关; (2)道德加工相关脑区构成的局部脑网络内部的功能连接强度(FC)能够有效预测框架效应强度。本研究首次利用静息态功能连接探讨个体的社会性决策受到语言表述方式影响的神经机制, 为进一步揭示社会决策中的个体差异提供了神经学证据。

关键词 社会性框架效应, 静息态功能磁共振, 功能连接, 右侧颞顶联合区, 道德脑网络
分类号 B849:C91; B845

1 引言

框架效应(Framing effect)指当以不同的方式描述同一个问题时, 决策者的选择偏好会受到描述方式的影响而改变。框架效应的强弱由问题的提出形式、社会规范以及决策者的人格特征等因素共同决定(Tversky & Kahneman, 1981)。根据个体的决策是否会影响到其他人, 框架效应可以分为非社会性和社会性两种, 分别对应于非社会性和社会性决策。经济决策中的增益/损失(Gain/Loss)框架效应是一个经典的非社会性框架：增益框架下, 个体会更加厌恶风险, 偏好确定的增益; 损失框架下, 个体会更加追求风险, 而回避确定的损失(de Martino et al., 2006)。在进行类似的非社会性决策时, 个体致力于最大化个人效益。此时, 框架通过调控个体对选项是否有利(或者是否具有高风险性)的判断来影响其选择。当个体的决策会影响到他人时(即进行社会性决策时), 个体会将他人的福祉、需求、态度以及社会评价等因素纳入考量。此时, 框架通过调控决策者对社会规范(如道德、公平、利他等)的偏好来影响其选择(Liu, Gu et al., 2020)。例如, 当研究者把一项社会性决策任务命名为“社区游戏”时, 与“华尔街游戏”这个名称相比, 被试的合作程度会显著提高(Liberman et al., 2004)。可见, 社会性框架效应与非社会性框架效应的心理机制是不一样的(Rege & Telle, 2004)。

近年来, 大量神经科学研究致力于揭示框架效应这一重要心理现象的脑机制。已有研究主要集中在非社会性框架上, 大多采用 Gain/Loss 范式。这些研究的主要发现有二：第一, 杏仁核(amygdala)和前扣带回(anterior cingulate cortex, ACC)是非社

收稿日期: 2020-04-17
* 国家自然科学基金(31900779, 31871109, 32071083, 32020103008)资助。
通信作者: 刘洁, E-mail: ljier06@gmail.com

会性框架效应的两个关键脑区；第二，特质焦虑水平越高的个体，表现出的框架效应越强(de Martino et al., 2006; Xu et al., 2013)。关于社会性框架效应的研究相对较少。考虑到二者的心理机制差异甚大，其神经机制也应有所不同。我们设计了一种新的任务范式：要求被试在个人经济利益和他人感受之间作出权衡，如果个体倾向于最大化个人利益，那么与他/她配对完成实验的另一个被试就会有更大的几率接受中等疼痛强度的电刺激。范式包含两种框架，一种是“伤害框架”：选择最大化个人利益而令他人遭受疼痛电击被描述为“主动伤害他人”；另一种是“帮助框架”：最大化个人利益而令他人遭受疼痛电击的选项被描述为“不帮助他人”。Liu 等已发表的研究指出，该范式可以有效诱发框架效应：相比于“帮助框架”，被试在“伤害框架”下更倾向于做出利他选择(即减少个人利益，避免他人疼痛)。功能磁共振结果表明：右侧颞顶联合区(right temporo-parietal junction, 右侧 TPJ)在该框架效应中起关键作用(Liu, Gu et al., 2020)。

尽管本质相同，但做出主动“伤害”的决定比被动地选择“不帮助”更容易被认为是不道德的。把“被试保留金钱，让另一被试接受疼痛电击”这一结果描述为伤害他人同时自己获利，强调了对他人“造成伤害”的后果，从而增强了道德信念与物质利益之间的冲突，引发个体回避这种冲突的动机。因此，在伤害框架下，相比于帮助框架，被试更倾向于选择牺牲自己的经济利益来避免伤害他人(Liu, Gu et al., 2020)。过往研究表明，右侧 TPJ 负责处理决策过程中与道德有关的冲突(Berns et al., 2012; Morishima et al., 2012)。因此这一脑区激活水平的变化，可能反映了不同框架下道德冲突的差异(Liu, Gu et al., 2020)。判断一个行为是否道德，通常取决于我们对其行为意图及结果的认知和评估(Young, Cushman et al., 2007; Young & Dungan, 2012)。研究发现在道德判断任务中，当被试加工行为意图时，右侧 TPJ 会被显著激活(Young & Koenigs, 2007; Young et al., 2011)。此外，右侧 TPJ 还涉及对他人意图的推测(即心理理论功能)和结果的信息整合加工(Koster-Hale et al., 2013; Young & Saxe, 2008)。先前研究发现，非社会性框架效应与个体的焦虑特质相关(Xu et al., 2013)。社会性框架效应也很可能与个体的道德相关特质有关。

静息态功能磁共振成像不受实验任务的限制和制约，能可靠地测量到大脑内在的自发活动，适用于探索反映个体差异的神经指标(Zou et al., 2008; 江琦 等, 2018)。静息态功能连接的方法也被普遍用来探索脑区之间的联系以及其与特定心理现象或行为之间的关系(Takeuchi et al., 2012; Zeng et al., 2012; 江琦 等, 2018)。静息态时的功能连接模式与任务态时的激活模式相似度近 80% (Biswal et al., 1995; Cole et al., 2016)。值得注意的是，静息态功能连接不仅被用来探索感知觉(Lewis et al., 2009; Taubert et al., 2011)，注意等认知功能上的个体差异(Mennes et al., 2010; Posner et al., 2014)，还被用于探索复杂社会决策中的个体差异。例如，Wang 等发现壳核(putamen)、TPJ 和内侧前额叶(medial prefrontal cortex, MPFC)静息态时的功能连接强度与利他惩罚中的责任分散效应相关(Wang et al., 2017)。之前对任务态数据的分析揭示了在被试内水平上右侧 TPJ 与社会性框架效应的相关性(Liu, Gu et al, 2020)。采用同批被试在任务前采集的静息态数据，可以进一步在被试间水平上，从大脑内在功能结构的角度探讨社会性框架效应的神经基础。

本研究选取低频振荡振幅(Amplitude of Low Frequency Fluctuation, ALFF) (Zou et al., 2008)探讨大脑内在活动与社会框架效应之间的相关关系，并通过脑网络功能连接(Functional Connectivity, FC)预测社会框架效应。先前研究发现灰质在静息态下的低频振荡信号振幅 ALFF 高于白质(Biswal et al., 1995)，且在对应认知过程的关键区域，患有精神疾病的个体也被发现 ALFF 异常(Zang et al., 2007; Hoptman et al., 2010)。健康个体在静息状态下的 ALFF 与任务诱发的 BOLD 信号以及个体的行为反应指标相关(Mennes et al., 2011)。这些发现表明，ALFF 在一定程度上反映了大脑静息状态下的局部自发活动(Zang et al., 2007)。其次，大量研究发现在同一神经解剖或功能系统内，不同区域之间的低频信号在时间波动上存在相关性(Biswal et al., 1995; Fox et al., 2005)；且不同脑区之间的功能连接强度及连接模式与健康个体的认知能力(Lewis et al., 2009)及社会性特征(Nostro et al., 2018; Wang et al., 2019)存在显著相关。综上可知，ALFF 和 FC 都是能够反映人类认知特点的脑信号度量。ALFF 反映静息状态下大脑局部的低频信号特征，FC 则反映不同脑区之间的功能整合。已有多项研究将这两种测量指标结合使用(如 de Vos et al., 2018; Wang et al., 2019)。因此，本研究拟采用 ALFF 分析方法探测与社会性框架效应相关的脑区，并进一步进行基于种

子点的 FC 分析, 以期全面阐述社会性框架效应的脑机制。

本研究将道德决策任务态功能磁共振研究的元分析中得到的激活脑区作为兴趣区域(包括 MPFC、尾状核[caudate]、颞上回[superior temporal gyrus, STG]、颞中回[Middle temporal gyrus, MTG]、额下回[Inferior frontal gyrus, IFG]等) (Garrigan et al., 2016), 建立局部脑网络, 与行为框架效应进行相关分析。并通过机器学习的方法进行多变量功能连接的相关分析, 从而对行为层面上的社会性框架效应进行预测。

2　方法

2.1　被试

30 名健康成年人被试(均为在校大学生)通过广告招募参加研究。所有人均为右利手, 视力或矫正视力正常。扫描期间头部旋转 > 1°或平移 > 1 mm 的 6 名被试数据被剔除, 24 名被试纳入统计分析(女性 12 名, 年龄 20.58 ± 1.91 岁[$M \pm SD$])。本研究根据《赫尔辛基宣言》的道德准则和原则进行, 经深圳大学医学院医学伦理委员会批准, 所有被试均签署知情同意书。

2.2　行为任务

本范式根据“趋避冲突” (Approach-Avoidance Conflict Task)任务改编而成, 该任务已被证明适用于功能磁共振成像研究(Aupperle et al., 2015; Aupperle et al., 2011)。核心逻辑来自 Crockett 等人(2014)提出的观点: “避免给他人造成痛苦是一种强大的行为动机”, 本范式将在自己的经济收益和他人的痛苦之间制造趋避困境。

本实验为单因素被试内设计。自变量为“框架”, 包含帮助框架与伤害框架两个水平。实验流程如下: 到达实验室后, 每个被试都被介绍给另一被试(由同性别的假被试扮演)。两名被试抽签决定各自在实验中的角色。通过操纵抽签结果, 真被试始终扮演“决策者”角色, 假被试则扮演“接受者”角色。具体操纵方式为: 两张不同颜色纸牌背面朝上摆在桌上, 一张正面写着“决策者”, 另一张写着“接受者”; 主试装作随意地邀请假被试先选牌, 此时假被试将选择事先约定好的牌(即写着“接受者”的牌); 真被试总会得到“决策者”的牌。实验后将要求被试对感受到的真实度进行评价, 以确保其相信情境设置。

被试被告知, 实验中接受者将受到一定次数的中等疼痛强度电击, 每次电击持续 50 ms, 电流强度为接受者在 0~9 疼痛量表上评分为 5 时所对应的强度。决策者将通过按键进行决策, 其决策结果将影响接受者在该试次中接受电击及决策者自己被扣除 5 元钱的概率。真被试随后被带到功能磁共振成像实验环境中, 且被告知“接受者”将在另一个房间通过网络连接参与任务。

每个试次中, 屏幕的左右两端呈现不同的图文, 分别表示两种可能的结果。在伤害框架条件下, 一侧显示的结果是“电击接受者(且不扣钱)”, 另一侧显示的结果是“不电击接受者并从你的报酬扣除 5 元”; 帮助框架条件下, 一侧结果是“帮助接受者免除电击并从你的报酬中扣除 5 元”, 另一侧结果是“不帮助接受者免除电击(且不扣钱)”。被试通过按键移动选择轴上的红色滑块来表达自己对于呈现在选择轴两侧结果的偏好。滑块的最终位置将决定两种结果在本试次中发生的概率。滑块越靠近某一侧, 处于该侧的结果发生概率就越大, 另一侧的结果发生概率相应减小。例如, 如果滑块放置于左侧或右侧的顶端, 则该侧结果发生的概率为 90%, 对侧结果发生的概率为 10%。如果被试选择将滑块置于中点, 则两侧的结果随机发生(各 50%) (图 1A)。两种结果的左右位置在试次间平衡, 滑块的初始位置在试次间随机。被试有 4 秒时间通过反应盒上的两个按键移动滑块位置, 并按第三个按键进行确认。每个试次中, “接受者”的照片呈现 2 秒钟。随后被试进入决策阶段, 经过 2~4 秒的等待后, 当前试次的结果将呈现 2 秒(图 1B)。每个试次的最终结果由计算机根据被试选择的概率生成。

值得注意的是, 决策者在这两种框架下所面对的是本质相同的选择, 即是否花费自己的 5 元钱来让他人免受电击。两种框架间唯一的区别在于对结果的描述方式—— “决策者不被扣钱, 接受者受到疼痛电击”在帮助框架下被描述为不帮助他人, 在伤害框架下则被描述为故意伤害他人; 另一种结果“决策者被扣除 5 元, 接受者不受疼痛电击”, 在帮助框架下被描述为帮助他人, 在伤害框架下则被描述为避免伤害他人。

实验前告知被试其报酬在 150~200 元之间, 实际获得的金额将与其在任务中的决策有关。实验后通过小程序随机抽取 10 个试次, 从 200 元基础费用上上扣除对应的钱数并支付给被试。实验过程中, 被试并不知道哪些试次会被抽取执行。因此, 对于被试而言, 最佳策略是以同样的态度对待每个试次。实验共 80 个试次, 分为 2 个区组, 持续约 30

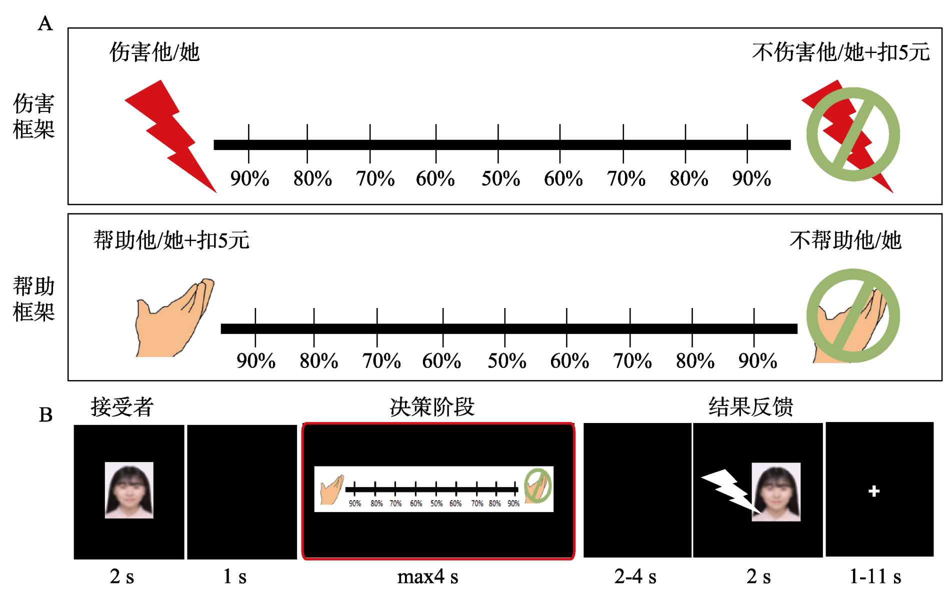


图 1 实验设计示意图 (A)伤害框架和帮助框架; (B)试次流程

分钟。每种条件(帮助/伤害框架)伪随机呈现。正式实验前，被试通过 8 个试次的练习熟悉任务。

2.3 fMRI 数据采集

使用西门子 3.0T 磁共振扫描仪进行数据采集。T1 加权结构像使用磁化准备快速采集梯度回波序列 (magnetization-prepared rapid acquisition gradient echo, MPRAGE)采集，具体扫描参数为：重复时间(repetition time, TR) = 2530 ms，回波时间(echo time, TE) = 2.96 ms，矩阵(matrix) = 256 × 256，体素大小(voxel size) = 1 mm × 1 mm × 1 mm，层数(slices) = 256，覆盖全脑。

被试先进行闭眼静息态扫描。使用多层 T2 加权回波平面成像序列(multiple slice T2-weighted echo planar imaging sequences, EPI)采集。具体扫描参数为：重复时间 = 2000 ms，回波时间= 30 ms，反转角(flip angle) = 90°，层厚(slice thickness) = 3.5 mm，矩阵= 64 × 64，体素大小= 3.5 mm × 3.5 mm × 4.2 mm，层数= 33。扫描时间为 6 分钟，共获得 180 个时间点的图像。静息态扫描结束后，经过 2 分钟的休息，进行任务态扫描。已发表论文(Liu, Gu et al., 2020)使用了任务态的核磁数据和行为数据。本论文使用了静息态数据和任务态扫描中收集的行为数据。二者的重叠部分仅在行为数据。

2.4 数据预处理

静息态数据使用 DPABI 软件(http://www.restfmri.net/forum/DPABI)进行预处理(Yan et al., 2016)。首先将原始 DICOM 数据转换为 NIFTI 格式，为去除磁共振信号起始状态不稳定和被试刚进入扫描仪时的不适应所带来的影响，删除前 10 个时间点的数据，剩余 170 个时间点进行后续处理。随后进行时间层校正(slice timing)，采用隔层扫描，共 33 层，参考层为全脑扫描过程中位于中间时间点的第 33 层。之后进行头动校正(realign)，采用 DARTEL 进行空间标准化(spatial normalization)，空间标准化后体素为 3 mm × 3 mm × 3 mm，将图像配准到标准 MNI (montreal neurological institute)空间。经过高斯平滑(smoothing) (平滑核[FWHW] = 6 mm)后，进行去信号线性漂移(detrend)，最后进行带通滤波(band filter)，滤波范围为 0.01 ~ 0.1 Hz。

2.5 数据分析

2.5.1 行为数据

被试在两种框架下，面对的是本质相同而描述方式不同的两难困境。我们预期不同框架会影响被试的选择偏好。框架效应被计算为两种框架下被试选择牺牲个人利益而令他人免受疼痛的倾向性(即助人倾向)的差异。具体来说，在伤害框架下，被试选择最靠近“不伤害他/她+扣除 5 元”一端选项时，赋权重值为 9，选项向另一端移动时权重递减。据此可知，被试选择最靠近“伤害他/她”一端选项时，赋权重值递减为 1。与之相应，在帮助框架下，被试选择最靠近“帮助他/她免除伤害+扣除 5 元”一端选项时，赋权重值为 9。先计算每个被试在两种框架下的平均助人度，再对两种框架下的助人度进行双边配对样本 t 检验。框架效应分数根据以下公式计算。该分数被作为社会性框架的行为指标，并纳入随后的静息态核磁数据分析。

框架效应分数 = 助人度$_{伤害框架}$ − 助人度$_{帮助框架}$

2.5.2　静息态 fMRI 数据

(1)体素水平分析。对预处理后的静息态图像，提取去除线性漂移后每个体素的时间序列，经过 0.01~0.1 Hz 带通滤波器提取滤波结果，经过快速傅立叶变化后得到功率谱，将功率谱开方后平均得到 ALFF。最后将 ALFF 除以全脑所有体素的平均 ALFF，得到标准化的 ALFF (mALFF)。为探讨静息状态下的 mALFF 指标与社会性框架效应的关系，采用 RESTplusV1.22 工具包(Jia et al., 2019)计算个体 mALFF 与框架效应分数之间的 *Pearson* 相关系数，并将被试的平均头动参数(mean Framewise Displacement, meanFD_Jenkinson, (Jenkinson et al., 2002)作为协变量回归掉。设置体素水平 $p < 0.005$，团块水平 $p < 0.05$ 的 GRF 标准进行多重比较校正。

(2)兴趣脑区(Region of Interest, ROI)水平分析。通过 ROI 之间的功能连接来探讨道德加工脑网络在社会框架效应中的作用，采用单变量功能连接分析和基于机器学习的预测模型两种分析方式。

选取种子点并提取信号。在体素水平分析中得到的显著脑区很可能在社会性框架效应中起到重要作用，因此将该脑区的峰值点作为第一个种子点。再根据道德相关任务态 fMRI 研究元分析的结果(Garrigan et al., 2016)选取与道德评估和道德决策相关的 12 个区域作为种子点，共得到 13 个种子点(详见表 1)。利用 RESTplus 工具包下的功能连接工具，以这 13 个点的 MNI 坐标为中心，取半径 6 mm 的球体作 ROI。利用 RESTplus 工具包在每个被试的功能连接脑图中提取出这 13 个 ROI 的信号值，得到 13 × 13 的功能连接矩阵(ROI 之间两两连接值共 78 个)。矩阵中的每一条边代表每两个 ROI 之间的相关性。为了消除由噪声导致的弱相关将小于 0.2 的值替换为 0 (Liu et al., 2017) (其他阈限结果见 3.3 结果有效性验证)。

单变量功能连接分析。计算 ROI 之间功能连接值与 z 标准转化后的框架效应分数的 *Pearson* 相关系数，并进行 False discovery rate (FDR)多重比较矫正。

基于机器学习的预测模型分析(多变量功能连接分析)。选取之前计算得到的 78 个 ROI 间功能连接值作为自变量，z 标准化后的框架效应分数作为因变量，使用机器学习方法建立预测模型。相关向量回归(relevance vector regression analysis, RVR)在脑−行为认知映射中表现出很高的预测性能(Cui & Gong, 2018)。因此，选择 RVR 建立相关向量回归模型。RVR 是一种在完全概率贝叶斯框架下的稀疏核的机器学习多元回归方法，即在贝叶斯框架下，在模型权重上引入零均值高斯先验，由一组超参数控制，每一个权重对应一个超参数。根据训练数据迭代估计这些超参数的最可能值。由于许多权重的后验分布在 0 附近急剧到达峰值，从而获得稀疏性，与非零权重相关的训练向量称为“相关”向量。随后通过计算预测分布，对于隐藏的输入向量(即 78 个功能连接值)使用权重上的优化后验分布来预测因变量(即框架效应分数)。本研究中，输入向量(特征)是标准化后的训练数据，通过留一法进行交叉验证获得模型的概化性估计，模型估计结果参数用 *Pearson* 相关系数表示。RVR 符合小样本统计学习与预测的贝叶斯机器学习理论，不包含特定于算法的参数，因此不需额外的计算资源用于参数估计(Tipping, 1999)。

RVR 中只有一些被称为“相关向量”的样本(小于训练样本量)用于拟合模型：$y(x)=\sum_{i=1}^{m} w_i\tau_i + \in$。其中 τ_i 是基函数，ϵ是均值等于 0 和方差等于 β 的正态分布。RVR 使用训练数据来构建回归模型：$y=\theta\omega+\epsilon$。其中，$y=[y_1,y_2,\cdots,y_n]^T$，$\theta=[\theta_1,\theta_2,\cdots,\theta_n]$，$\theta_i=[\tau_i(x_1),\tau_i(x_2),\cdots,\tau_i(x_n)]^T$。每个矢量 θ_i 包括输入向量的基函数 τ_i 的值，且是相关向量。通过使用条件分布 $p(y|\alpha,\beta)=N(y|0, C)$的最大似然估计来确定模型参数 β。其中，$C=\beta I_n+\phi A^{-1}\phi^T$。为了生成适合于相关向量模型的稀疏回归模型，我们假定 w_i 和 β^{-1} 的先验分布，如 $p(w_i|\alpha_i)=N(0, \alpha_i^{-1})$，形成了同样的相关向量机构造。在找到参数 α_i 和 β 的最佳点估计后，由公式 $w=\beta(\beta\phi^T\phi+A)^{-1}\phi y$ 计算出权重向量的点估计值 w。

我们进一步采用置换检验(Permutation test)考察这一参数的显著性。使用留一交叉验证法，将预测因变量随机分配，让 RVR 模型重新运行 1000 次，即获得模型参数随机分布，该分布反映了模型没有超过偶然性的零假设。之后将置换值大于或等于真值的次数除以 1000，即为所得到模型参数(即回归系数)的 P 值。

$$P=\frac{\text{置换检验的次数}<\text{实际精度}+1}{\text{置换检验的次数}+1}$$

模型计算得出的特征权重指示回归模型中每个特征的重要性。绝对权重越大代表该连接在模型中预测性越强。在线性 RVR 模型中，权重为正的连接表示该连接越大，模型预测性越好；权重为负的连接表示该连接越小，模型预测性越好。

3 结果

3.1 行为结果

双边配对样本 t 检验发现, 伤害框架下的助人度显著高于帮助框架下(3.60 ± 1.82 vs. 3.06 ± 1.72 [平均值±标准差], $t(23) = 3.81$, $p = 0.001$, Cohen' $d = 0.31$) (图 2A)。这一结果表明不同框架影响了个体的选择倾向。双边单样本 t 检验表明框架效应分数显著大于 0, $t(23) = 4.12$, $p < 0.001$ (图 2B)。对反应时进行双边配对样本 t 检验, 发现两种框架下被试的决策时间没有显著差异($p = 0.86$)。

3.2 脑成像数据结果

3.2.1 体素水平分析

为了探索社会性框架效应与大脑静息状态活动之间的关联, 首先计算被试大脑中 ALFF 值与框架效应分数的 *Pearson* 相关系数。发现右侧 TPJ (peak MNI 坐标[39, −54, 6]; 60 个体素) 的 mALFF 与框架效应分数呈正相关, $R_{peak} = 0.74$ (体素水平 $p < 0.005$, 团块水平 $p < 0.05$, GRF 矫正) (图 3A)。

3.2.2 基于 ROI 水平分析的功能连接强度

所有 ROI 的种子点见表 1。其中 ROI1 来自体素水平分析中 mALFF 与框架效应分数显著相关的峰值坐标(即右侧 TPJ), 其余 12 个 ROI 来自元分析(Garrigan et al., 2016)所得到的道德脑网络。

单变量功能连接相关分析得出: 左侧内侧眶额皮层(Orbital medial prefrontal cortex, omPFC) (MNI [−4, 48, −6])和左侧尾状核(MNI [−12, 4, 12])之间的功能连接值与框架效应分数显著相关($r = 0.66$, $p < 0.001$, 通过 FDR 矫正) (图 4)。

机器学习预测模型分析发现, 基于 13 个 ROI 间功能连接值的 RVR 模型可以显著预测被试的框架效应分数($r = 0.42$, p=0.035)。绝对权重从大到小的前 13 个功能连接($|w| > 0.15$)分别为, 左侧 MPFC (ROI7)与左侧尾状核(ROI13)的连接 ($w = 0.35$); 左侧 MPFC (ROI2)与左侧尾状核 (ROI13)的连接 ($w = 0.27$); 左侧 MPFC 内部亚区之间的连接 (ROI2 与 ROI7, $w = 0.24$), 左侧扣带回(Cingulate gyrus, CG) (ROI4)与右侧 MPFC (ROI6)的连接($w = −0.24$), 左侧 MPFC (ROI2)与右侧 IFG (ROI12)的连接($w = −0.23$), 右侧 MPFC 内部亚区之间的连接(ROI6 与 ROI10, $w = −0.23$), 右侧 STG (ROI5)与右侧 IFG (ROI12)的连接($w = −0.21$), 右侧 STG (ROI5)与右侧 MPFC (ROI10)的连接($w = −0.21$), 右侧 MPFC

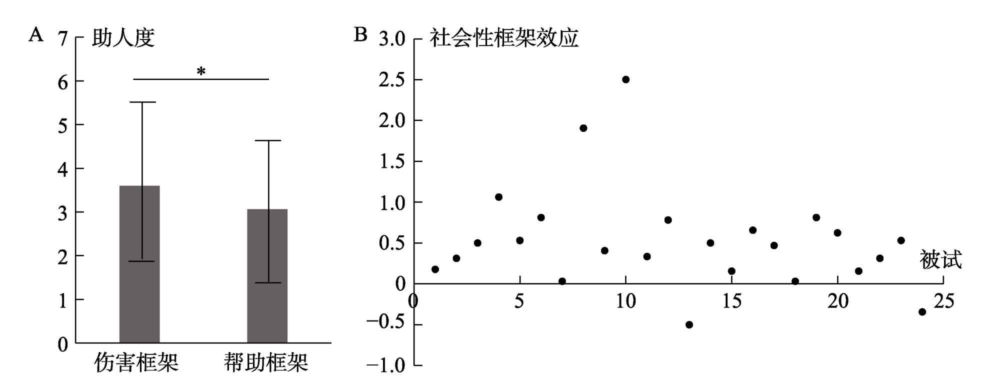


图 2 行为结果(A)伤害框架和帮助框架下的助人度; (B)所有被试的框架效应分数。

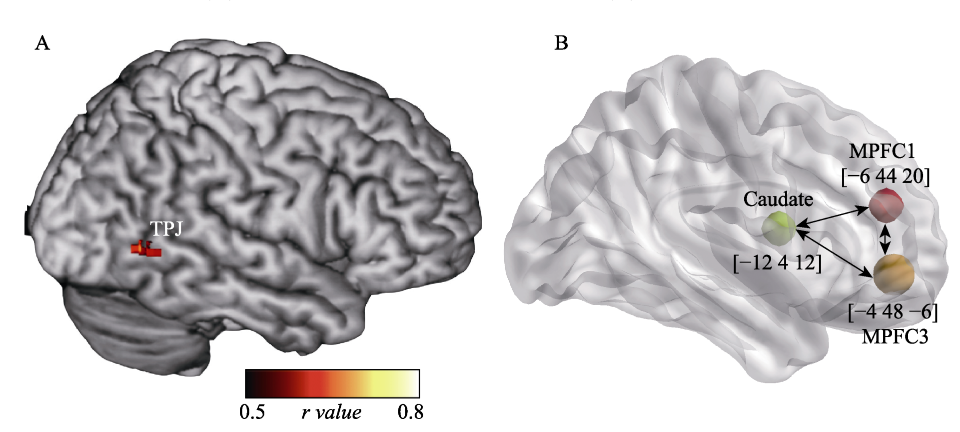


图 3 (A)体素水平分析结果; (B)左侧 MPFC1、MPFC3、尾状核之间的功能连接值与框架效应分数显著相关(不同颜色的球体表示对应 ROI; 黑色双箭头表示功能连接; 仅显示绝对权重排在前三位的连接)

表 1　种子点信息

来源	编号	x	y	z	半球	脑区
体素水平分析结果	ROI1	39	−54	6	右	颞顶联合区(TPJ)
道德脑网络(元分析)	ROI 2	−6	44	20	左	内侧前额叶(MPFC1)
	ROI 3	−44	−56	18	左	颞上回(STG1)
	ROI 4	−2	−56	26	左	扣带回(CG)
	ROI 5	50	6	−20	右	颞上回(STG2)
	ROI 6	2	54	2	右	内侧前额叶(MPFC2)
	ROI 7	−4	48	−6	左	内侧前额叶(MPFC3)
	ROI 8	−44	−64	20	左	颞中回(MTG1)
	ROI 9	−2	−60	30	左	楔前叶(Precuneus)
	ROI 10	2	44	36	右	内侧前额叶(MPFC4)
	ROI 11	44	−60	24	右	颞中回(MTG)
	ROI 12	36	28	−12	右	额下回(IFG)
	ROI 13	−12	4	12	左	尾状核(Caudate)

注：TPJ = Temporo-parietal junction; MPFC = Medial prefrontal cortex; STG= Superior temporal gyrus; CG=Cingulate gyrus; MTG= Middle temporal gyrus; IFG= Inferior frontal gyrus.

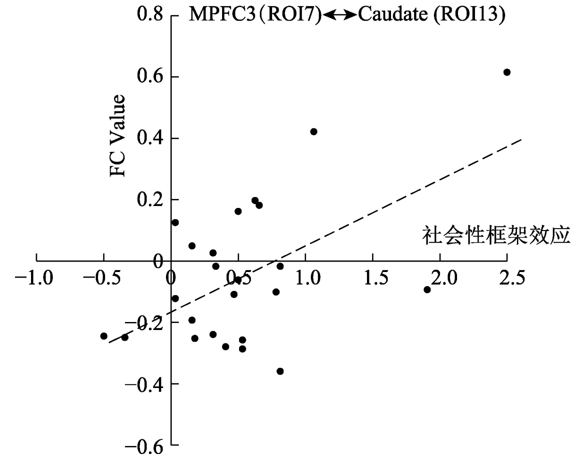


图 4　内侧前额叶皮层(MPFC3)和尾状核(Caudate)之间功能连接与社会性框架效应分数的相关点图

(ROI6)与左侧尾状核(ROI13)的连接(w = 0.20), 左侧 CG (ROI4)与右侧 MTG (ROI11)的连接(w = −0.19), 左侧 CG (ROI4)与左侧 MTG (ROI8)的连接(w = 0.18), 左侧 MTG (ROI8)的连接与右侧 IFG (ROI12)的连接(w = −0.18), 右侧 MPFC (ROI6)与右侧 IFG (ROI12)的连接(w = 0.17), 左侧 MPFC (ROI7)及右侧 MPFC (ROI10)的连接(w = −0.15), 左侧 MTG (ROI8)和左侧 Precuneus (ROI9)的连接(w = 0.15) (图 3B)。

3.3　结果有效性验证

(1)通过对预处理图像执行“scrubbing”程序(Power et al., 2012)进一步验证了头部运动对功能连接的影响。删除了逐帧位移(framewise displacement, FD)超过 0.2 mm 阈值的时间点及其相邻的时间点(前 2 个和后 1 个)。由于 Scrubbing 程序会删除信号, 无法进行傅里叶变换计算 ALFF 值, 从而导致无法进行体素水平的相关分析。对于 ROI 水平的分析, 我们使用删除后的数据提取 13 个 ROI 的功能连接矩阵, 计算静息态功能连接与框架效应分数的相关。发现这一结果与原有结果模式一致：单变量功能连接的相关分析得出左侧 MPFC (ROI7, MNI [−4, 48, −6])和左侧尾状核(ROI13, MNI [−12, 4, 12])之间的功能连接值与框架效应分数显著相关(r = 0.69, $p < 0.001$, 通过 FDR 矫正)。重新建立多变量的预测模型, 预测分数和框架效应分数之间的相关性略有变化, 但仍然显著($r = 0.41$, $p = 0.045$)。

(2)鉴于预处理步骤中的空间平滑化可能会在体素之间引入人为的局部相关性, 而这些相关性与其功能连接无关, 因此我们使用平滑核(FWHW) = 4 mm 的参数, 来验证主要结果。体素水平分析得出, mALFF 与框架效应相关的脑区在右侧 TPJ (peak MNI 坐标[39, −54, 6]; 41 个体素), $R_{peak} = 0.75$ (体素水平 $p < 0.00$, 团块水平 $p < 0.05$, 通过 GRF 矫正)。计算与框架效应相关的静息态功能连接发现, 左侧 MPFC (ROI7)与左侧尾状核 (ROI13)的连接与框架效应相关, 通过 FDR 矫正(r = 0.66, p < 0.001)。重新建立多变量预测模型, 预测分数和框架效应分数之间的相关性会略有变化, 但仍然显著($r = 0.44$, $p = 0.025$)。

(3)去除全脑信号。排除头动等 24 个全脑信号影响后, 重新分析结果发现, 在体素水平 mALFF 与框架效应分数相关的脑区为右侧 TPJ (peak MNI 坐标[39, −57, 3]; 72 个体素), $R_{peak} = 0.68$ ($p < 0.01$, 未通过 GRF 矫正)。在 ROI 水平, 单变量功能连接相关分析得出左侧 MPFC (ROI7)和尾状核 (ROI7)之间的功能连接值与框架效应显著分数相关(r = 0.54, p = 0.007)。机器学习预测模型分析发现, 基于 13 个 ROI 间功能连接值的 RVR 模型可以边缘显著地预测框架效应分数($r = 0.27$, p=0.091)。

(4)选取不同功能连接阈限值。将功能连接阈值设定为不同数值(包括 0.1、0.3、0.4、0.5)并重复其后的统计分析。结果如下：(1)阈限设定为 0.1 时：单变量功能连接相关分析发现左侧 MPFC (ROI7)和尾状核(ROI7)之间的功能连接值与框架效应分数显著相关($r = 0.63$, $p = 0.001$)。机器学习预测模型分析发现, 基于 13 个 ROI 间功能连接值的 RVR 模型可以边缘显著地预测被试的框架效应分数(r =

0.34, $p = 0.064$); (2)阈限设定为 0.3 时：单变量功能连接相关分析发现左侧 MPFC (ROI7)和左侧尾状核(ROI13)之间的功能连接值与框架效应分数显著相关($r = 0.66$, $p < 0.001$)。机器学习预测模型分析发现, RVR 模型可以边缘显著地预测被试的框架效应得分($r = 0.31$, $p = 0.078$); (3)当阈限设定为 0.4 和 0.5 时，由于超过 80%的数据点未到达相关系数阈限被置为 0，现有分析无法正确反映实际数据情况。因此，在这两个阈限值下无法进行相关分析。据此可见，当 ROI 之间相关系数取值变化在一定范围内时，结果模式相对稳定。

4 讨论

本研究为社会性决策受到语言表述方式的影响提供了证据，揭示了社会性框架效应的神经机制。通过单变量的功能连接强度和基于机器学习的多变量连接模式分析，发现右侧 TPJ 的 ALFF 值、道德加工脑网络中左侧 MPFC 与尾状核之间的功能连接强度及该网络内部的功能连接模式可以有效预测行为层面上社会性框架效应的强度。

与任务态结果一致，静息态分析结果再次凸显了右侧 TPJ 在社会性框架效应中的关键作用。静息态发现的 ALFF 指标与框架效应强度相关的脑区和同批被试的任务态中发现的关键脑区都定位在右侧 TPJ (Liu, Gu et al., 2020)。ALFF 指标可以在一定水平上反映静息状态下神经元自发活动水平的高低，也可以有效反映认知神经活动。例如，任务态研究发现认知障碍人群中，其认知障碍对应的关键脑区在静息态中也表现出异常的 ALFF 值(Zang et al., 2007)，而健康被试在静息态下的 ALFF 值与任务态诱发的 BOLD 值及被试的行为指标之间存在相关(Mennes et al., 2011; Wei et al., 2012)。近年来大量研究揭示了右侧 TPJ 在社会认知和社会决策中的核心作用(Krall et al., 2015)。例如，有研究发现该脑区的激活强度反映了决策过程的冲突水平(Berns et al., 2012; Morishima et al., 2012)。本范式中，相比于帮助框架，伤害框架可能会诱发更强的道德冲突感，因为相比于“不帮助”，选择“去伤害”会更加强烈地违反社会规范和道德标准。因此，右侧 TPJ 在静息态下更强的自发活动(即更高的 ALFF 值)可能反映了被试对道德冲突的敏感程度。此外，右侧 TPJ 在心理理论，特别是道德意图加工方面也起到重要作用。前人研究发现，该脑区在道德判断中起到了整合行为意图及结果的作用，即不仅考虑行为所导致的后果，还会结合行为人的意图和动机(例如，是否有意伤害)来进行道德评价(Young & Saxe, 2008; 甘甜 等, 2013; 罗俊 等, 2017)。相比于帮助框架，伤害框架也包含有意伤害的成分，因而更容易在道德判断中被认为是“不道德的，不可接受的”。因此，右侧 TPJ 的自发活动也可能影响个体在道德判断和决策中对意图和可能发生的后果的整合。

本研究中，社会性框架效应发生的基础是在某一框架下(如伤害框架)强调选项的道德性，而在另一框架下不强调，并由此在被试身上引发对同质决策困境的不同倾向。因此，道德加工脑网络可能支撑了这一行为效应的发生。我们从道德相关 fMRI 任务态的元分析中提取了在道德评估和道德决策两类任务中一致性被激活的 12 个脑区(Garrigan et al., 2016)，并采用多种方式探索和验证了由这些 ROI 构成的道德脑网络与社会性框架效应之间的关系。结果发现，由道德相关任务态元分析发现的脑区所构成的局部脑网络可以有效预测框架效应的强度。

先前研究发现，基底核区域的神经元对预期的道德水平敏感，在道德判断中，如果当前情境违背了个体预期的道德水平(例如，目睹好人做了不道德的事)，该区域的神经元会有更强的激活(Decety & Porges, 2011; Shenhav & Greene, 2010; Yoder & Decety, 2014)。其中，尾状核在多种类型的道德决策任务中都被激活，例如捐赠行为、利他惩罚等。当观看伤害他人的行为时，精神病态(psychopath)组在该区域的神经活动比控制组显著降低(Yoder et al., 2015)。研究者据此提出，尾状核对道德相关的冲突敏感(Liu et al., 2019)。本范式中，尾状核可能与对道德冲突的觉察有关，反映了个体对道德冲突信息的敏感程度。MPFC 在经济决策和道德决策中都起到关键作用(Hare et al., 2009; Rolls et al., 2008)。研究表明，该区域可能反映整个决策过程中的价值评估，即通过解码决策过程中涉及的多个相关因素来生成整合性价值并据此进行决策(Kaanders & Hunt, 2018)。相关研究也证明了 MPFC 在道德决策中的类似作用，即对道德相关价值的评估和整合，该区域受损会直接影响到道德判断能力(Brosch et al., 2012; Shenhav & Greene, 2014)。此外，元分析表明 MPFC 和右侧 TPJ 一样，都是心理理论功能中重要的一环。因此，尾状核与 MPFC 之间的功能连接强度，可能反映了对道德冲突的觉察以及道德冲突在决策中的权重。具体来说，个体对道德

冲突敏感性越高，伤害框架相比于帮助框架下更强的道德冲突就越容易被觉察到，而二者之间的功能连接越强，这种道德冲突的差异就能更好地被纳入到最终的价值整合中，从而影响决策行为。值得注意的是，功能连接分析并不反应两个脑区之间的连接方向。因此，还有一种可能性是，MPFC 与心理理论能力(个体推测和体验他人想法、评价、态度的能力)有关。在当前的范式中，越能体会到他人(即接受者)被主动伤害时的不愉悦感(相比于只是“不被帮助”)，越容易体会到道德冲突感，从而影响决策行为。

需要指出的是，在另一篇道德相关任务态 fMRI 研究元分析中，右侧 TPJ 也被发现激活(Bzdok et al., 2012) 。而在本研究选择的这篇更近期发表的元分析中则没有发现右侧 TPJ 的激活。这可能与论文筛选标准或元分析的参数设置有关，但是这一点提醒我们不能将右侧 TPJ 从道德加工脑网络中完全地分离出去。

本研究存在一些不足。首先，被试量相对较少，近期研究表明机器学习的准确率和稳定性会随着样本量的增大而提升(Cui & Gong, 2018)。虽然有先前对特定行为指标进行静息态数据相关分析的研究也使用了类似规模的样本量(Liu, Yuan et al., 2020)，且本研究谨慎地对静息态数据采用了多角度的分析和验证，尽量减少无关变量的干扰，但未来研究还需要通过增加样本量进行更有力地证明；其次，被试在任务中被扣除的金钱并不是已经拥有的，而是尚未支付的，这一差异可能降低本实验的生态效度。此外，先前研究表明，道德判断和认知共情等特质存在性别差异。本研究中虽不存在性别差异导致的干扰(男性和女性数量相同)，然而由于被试量较小，难以进一步进行性别之间的比较。再次，本研究结果在去除全脑信号的验证性分析中有所削弱。考虑到以往研究中也发现去除全脑信号削弱功能连接与行为反应之间联系的情况(Murphy & Fox, 2017; Gotts et al., 2013)，我们认为本研究的结果是比较稳定的。最后，本研究采用的行为学指标仅来自于一项实验任务，未来仍需要来自其他范式的行为学指标对此结果进一步验证。

5　总结

本研究基于新颖的实验范式，利用静息态功能成像技术探讨了社会性框架效应的脑机制。结果发现，右侧 TPJ 的低频振荡振幅和道德加工局部脑网络的内部功能连接模式及强度可以有效地预测个体的社会性框架效应强度。本研究首次利用静息态功能连接探讨社会性决策在多大程度上受到表述方式(即框架的)影响，进一步揭示了社会性决策中个体差异的神经机制。

参　考　文　献

Aupperle, R. L., Melrose, A. J., Francisco, A., Paulus, M. P., & Stein, M. B. (2015). Neural substrates of approach-avoidance conflict decision-making. *Human Brain Mapping*, *36*(2), 449–462.

Aupperle, R. L., Sullivan, S., Melrose, A. J., Paulus, M. P., & Stein, M. B. (2011). A reverse translational approach to quantify approach-avoidance conflict in humans. *Behavioural Brain Research*, *225*(2), 455–463.

Berns, G. S., Bell, E., Capra, C. M., Prietula, M. J., Moore, S., Anderson, B., ... Atran, S. (2012). The price of your soul: Neural evidence for the non-utilitarian representation of sacred values. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1589), 754–762.

Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo‐planar MRI. *Magnetic Resonance in Medicine*, *34*(4), 537–541.

Brosch, T., Coppin, G., Schwartz, S., & Sander, D. (2012). The importance of actions and the worth of an object: Dissociable neural systems representing core value and economic value. *Social Cognitive and Affective Neuroscience*, *7*(5), 497–505.

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure Function*, *217*(4), 783–796.

Cole, M. W., Ito, T., Bassett, D. S., & Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nature Neuroscience*, *19*(12), 1718–1726.

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(48), 17320–17325.

Cui, Z., & Gong, G. (2018). The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *Neuroimage*, *178*, 622–637.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*(5787), 684–687.

De Vos, F., Koini, M., Schouten, T. M., Seiler, S., van der Grond, J., Lechner, A., ... Rombouts, S. A. R. B. (2018). A comprehensive analysis of resting state fMRI measures to classify individual patients with Alzheimer's disease. *Neuroimage*, *167*, 62–72.

Decety, J., & Porges, E. C. (2011). Imagining being the agent of actions that carry different moral consequences: An fMRI study. *Neuropsychologia*, *49*(11), 2994–3001.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(27), 9673–9678.

Gan, T., Shi, R., Liu, C., & Luo, Y. J. (2018). Cathodal transcranial direct current stimulation on the right temporo-parietal junction modulates the helpful intention processing. *Acta Psychologica Sinica*, *45*(9), 1004–1014.
[甘甜, 李万清, 唐红红, 陆夏平, 李小俚, 刘超, 罗跃嘉. (2013). 经颅直流电刺激右侧颞顶联合区对道德意图加工的影响. *心理学报*, *45*(9), 1004–1014.]
Garrigan, B., Adlam, A. L. R., & Langdon, P. E. (2016). The neural correlates of moral decision-making: A systematic review and meta-analysis of moral evaluations and response decision judgements. *Brain and Cognition*, *108*, 88–97.
Gotts, S. J., Saad, Z. S., Jo, H. J., Wallace, G. L., Cox, R. W., & Martin, A. (2013). The perils of global signal regression for group comparisons: A case study of Autism Spectrum Disorders. *Frontiers in Human Neuroscience*, *7*, 356.
Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*, *324*(5927), 646–648.
Hoptman, M. J., Zuo, X. N., Butler, P. D., Javitt, D. C., D'Angelo, D., Mauro, C. J., & Milham, M. P. (2010). Amplitude of low-frequency oscillations in schizophrenia: A resting state fMRI study. *Schizophrenia Research*, *117*(1), 13–20.
Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, *17*(2), 825–841.
Jia, X.-Z., Wang, J., Sun, H.-Y., Zhang, H., Liao, W., Wang, Z., ... Zang, Y.-F. (2019). RESTplus: An improved toolkit for resting-state functional magnetic resonance imaging data processing. *Science Bulletin*, *64*(14), 953–954.
Jiang, Q., Hou, L., Qiu, J., Li, C., & Wang, H. (2018). The relationship between the caudate nucleus-orbitomedial prefrontal cortex connectivity and reactive aggression: A resting-state fMRI study. *Acta Psychologica Sinica*, *50*(6), 655–666.
[江琦, 侯璐璐, 邱江, 李长燃, 王焕贞. (2018). 尾状核−眶部内侧前额叶的功能连接与反应性攻击的关系: 基于静息态功能磁共振研究. *心理学报*, *50*(6), 655–666.]
Kaanders, P., & Hunt, L. T. (2018). Neuroscience: Intracranial recordings of value. *Current Biology*, *28*(18), 1106–1108.
Koster-Hale, J., Saxe, R., Dungan, J., & Young, L. L. (2013). Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(14), 5648–5653.
Krall, S. C., Rottschy, C., Oberwelland, E., Bzdok, D., Fox, P. T., Eickhoff, S. B., ... Konrad, K. (2015). The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. *Brain Structure Function*, *220*, 587–604.
Lewis, C. M., Baldassarre, A., Committeri, G., Romani, G. L., & Corbetta, M. (2009). Learning sculpts the spontaneous activity of the resting human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(41), 17558–17563.
Liberman, V., Samuels, S. M., & Ross, L. (2004). The Name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. *Personality & Social Psychology Bulletin*, *30*(9), 1175–1185.
Liu, J., Gu, R., Liao, C., Lu, J., Fang, Y., Xu, P., ... Cui, F. (2020). The Neural mechanism of the social framing effect: Evidence from fMRI and tDCS studies. *Journal of Neuroscience*, *40*(18), 3646–3656.
Liu, J., Liao, C., Lu, J., Luo, Y.-J., & Cui, F. (2019). Moral contagion: Devaluation effect of immorality on hypothetical judgments of economic value. *Human Brain Mapping*, *40*(7), 2076–2088.
Liu, J., Xia, M., Dai, Z., Wang, X., Liao, X., Bi, Y., & Yong, H. (2017). Intrinsic brain hub connectivity underlies individual differences in spatial working memory. *Cerebral Cortex*, *27*(12), 5496–5508.
Liu, J., Yuan, B., Luo, Y.-J., & Cui, F. (2020). Intrinsic functional connectivity of medial prefrontal cortex predicts the individual moral bias in economic valuation partially through the moral sensitivity trait. *Brain Imaging and Behavior, 14*(5), 2024–2036.
Luo, J., Ye, H., Zheng, H., Jia, Y., Chen, Z., & Huang, D. (2017). Modulating the activities of right and left temporo-parietal junction influences the capability of moral intention processing: A transcranial direct current stimulation study. *Acta Psychologica Sinica*, *49*(2), 228–240.
[罗俊, 叶航, 郑昊力, 贾拥民, 陈姝, 黄达强. (2017). 左右侧颞顶联合区对道德意图信息加工能力的共同作用——基于经颅直流电刺激技术. *心理学报*, *49*(2), 228–240.]
Mennes, M., Kelly, C., Zuo, X-N., Di Martino, A., Biswal, B.B., Castellanos, F.X., & Milham, M.P. (2010). Inter-individual differences in resting-state functional connectivity predict task-induced BOLD activity. *Neuroimage*, *50*(4), 1690–1701.
Mennes, M., Zuo, X-N., Kelly, C., Di Martino, A., Zang, Y. F., Biswal, B., … Milham, M. P. (2011). Linking inter-individual differences in neural activation and behavior to intrinsic brain dynamics. *Neuroimage*, *54*(4), 2950–2959.
Morishima, Y., Schunk, D., Bruhin, A., Ruff, C. C., & Fehr, E. (2012). Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*, *75*(1), 73–79.
Murphy, K., & Fox, M. D. (2017). Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage*, *154*, 169–173.
Nostro, A. D., Müller, V. I., Varikuti, D. P., Pl?Schke, R. N., Hoffstaedter, F., Langner, R., ... Eickhoff, S. B. (2018). Predicting personality from network-based resting-state functional connectivity. *Brain Structure & Function. 223*, 2699–2719
Posner, J., Park, C., & Wang, Z. (2014). Connecting the dots: A review of resting connectivity MRI studies in attention-deficit/hyperactivity disorder. *Neuropsychology Review*, *24*(1), 3–15.
Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage, 59*(3), 2142–2154.
Rege, M., & Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics*, *88*(7-8), 1625–1644.
Rolls, E. T., McCabe, C., & Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex, 18*(3), 652–663.
Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, *67*(4), 667–677.
Shenhav, A., & Greene, J. D. (2014). Integrative moral judgment: Dissociating the roles of the amygdala and ventromedial prefrontal cortex. *Journal of Neuroscience, 34*(13), 4741–4749.
Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T.,

Nouchi, R., & Kawashima, R. (2012). The association between resting functional connectivity and creativity. *Cerebral Cortex, 22*(12), 2921−2929.

Taubert, M., Lohmann, G., Margulies, D. S., Villringer, A., & Ragert, P. (2011). Long-term effects of motor training on resting-state networks and underlying brain structure. *Neuroimage, 57*(4), 1492–1498.

Tipping, M. E. (1999). *The Relevance Vector Machine*. Paper presented at the Advances in Neural Information Processing Systems 12, [NIPS Conference, Denver, Colorado, USA, November 29 - December 4, 1999].

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.

Wang, L., Lu, X., Gu, R., Zhu, R., Xu, R., Broster, L. S., & Feng, C. (2017). Neural substrates of context - and person - dependent altruistic punishment. *Human Brain Mapping. 38*(11), 5535–5550.

Wang, Z., Wang, Y., Sweeney, J. A., Gong, Q., Lui, S., & Mosconi, M. W. (2019). Resting-state brain network dysfunctions associated with visuomotor impairments in autism spectrum disorder. *Frontiers in Integrative Neuroscience*, *13*, 17.

Wei, T., Liang, X., He, Y., Zang, Y., Han, Z., Caramazza, A., & Bi, Y. (2012). Predicting conceptual processing capacity from spontaneous neuronal activity of the left middle temporal gyrus. *Journal of Neuroscience, 32*(2), 481–489.

Xu, P., Gu, R., Broster, L. S., Wu, R., van Dam, N. T., Jiang, Y., ... Luo, Y. J. (2013). Neural basis of emotional decision making in trait anxiety. *Journal of Neuroscience, 33*(47), 18641–18653.

Yan, C. G., Wang, X. D., Zuo, X. N., & Zang, Y. F. (2016). DPABI: Data processing & analysis for (resting-state) brain imaging. *Neuroinformatics, 14*(3), 339–351.

Yoder, K. J., & Decety, J. (2014). The Good, the bad, and the just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others. *Journal of Neuroscience*, *34*(12), 4161–4166.

Yoder, K. J., Harenski, C., Kiehl, K. A., & Decety, J. (2015). Neural networks underlying implicit and explicit moral evaluations in psychopathy. *Translational Psychiatry*, *5*(8), e625.

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The Neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(20), 8235–8240.

Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, *7*(1), 1–10.

Young, L., & Koenigs, M. (2007). Investigating emotion in moral cognition: A review of evidence from functional neuroimaging and neuropsychology. *British Medical Bulletin, 84*(1), 69–79.

Young, L., & Saxe, R. (2008). The neural basis of belief encoding and integration in moral judgment. *Neuroimage, 40*(4), 1912–1920.

Young, L., Scholz, J., & Saxe, R. (2011). Neural evidence for "intuitive prosecution": The use of mental state information for negative moral verdicts. *Social Neuroscience, 6*(3), 302–315.

Zang, Y. F., He, Y., Zhu, C.-Z., Cao, Q.-J., Sui, M.-Q., Liang. M., ... Wang, Y. F. (2007). Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain and Development, 29*(2), 83–91.

Zeng, L.-L., Shen, H., Liu, L., Wang, L., Li, B., Fang, P., ... Hu, D. (2012). Identifying major depression using whole-brain functional connectivity: A multivariate pattern analysis. *Brain: A Journal of Neurology*, *135*(5), 1498–1507.

Zou, Q. H., Zhu, C.-Z., Yang, Y., Zuo, X.-N., Long, X.-Y., Cao, Q.-J., ... Zang, Y. F. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: Fractional ALFF. *Journal of Neuroscience Methods*, *172*(1), 137–141.

Functional connectivities of the right temporoparietal junction and moral network predict social framing effect: Evidence from resting-state fMRI

CUI Fang[1,2], YANG Jiamiao[1], GU Ruolei[3,4], LIU Jie[2,1]

([1] *School of Psychology, Shenzhen University, Shenzhen* 518060, *China*)

([2] *Center for Brain Disorders and Cognitive Neuroscience, Shenzhen University, Shenzhen* 518060, *China*)

([3] *CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing* 100101, *China*)

([4] *Department of Psychology, University of Chinese Academy of Sciences, Beijing* 100049, *China*)

Abstract

As an important cognitive bias, the framing effect shows that individuals' decision preferences are sensitive to the verbal description (i.e., frame) of options. The social framing effect could be distinguished from the non-social one according to whether the decision would influence others. The psychological mechanisms of the non-social framing effect (e.g., Gain/Loss framing effect) and that of the social one are essentially different. When making non-social decisions, frames affect people's judgment of which option is more beneficial or less risky. When making social decisions, frames affect people's preferences through the influence of other-regarding concerns and social norms.

In the present study, a new paradigm was developed to induce the social framing effect. We asked

participants to make a tradeoff between economic benefits and the feelings of others; when participants showed a stronger preference for income maximization, the probability for their partners to receive a painful electrical shock would increase proportionally. This decision was described as either a "harm" to, or simply "not helping" other persons in two frame conditions. 30 participants (age: 20.58 ± 1.91 years old) were enrolled in the experiment and 24 of them were included in the final analysis. The resting-state functional magnetic resonance (rs-fMRI) data was acquired using the Echo Planar Imaging (EPI) sequence from a 3-T Siemens scanner. This scanning acquired 180 volumes with TR = 2 s (lasting 6 min). Rs-fMRI data were processed and analyzed using the DPABI and RESTplus toolbox to calculate the amplitude of low-Frequency Fluctuation (ALFF) and Functional Connectivity (FC).

On the behavioral level, we found that participants made more prosocial decisions in the Harm frame compared to the Help frame condition, resulting in a significant social framing effect. For the resting fMRI analysis, we first run a whole-brain correlation analysis between ALFF and the behavioral index and found the ALFF of the right temporoparietal junction (rTPJ) could significantly predict the behavioral index of the social framing effect. Considering the observed social framing effect would result from different levels of moral conflict between Harm and Help frames, we predicted that it would be closely related to the moral network. Therefore, we further localized 12 seeds from a new, meta-analysis of functional MRI studies for moral processing. Seed-based FC analysis showed that the functional connectivity between the medial prefrontal cortex and the caudate was significantly associated with the behavioral index of the social framing effect. Multivariate machine learning-based regression analysis further confirmed these results, suggesting the importance of rTPJ and moral network for the observed social framing effect.

The present study is based on a novel experimental paradigm, using resting functional imaging techniques to explore the brain mechanism of the social framing effect. We found that the ALFF value of the right TPJ and the strength of the functional connectivity value between the medial prefrontal lobe and the caudate within a moral network can effectively predict the social framing effect. This study is the very first one to explore the extent to which individual social decision-making can be influenced by verbal description and its underlying neural mechanisms, which shed light on the further exploration of individual differences in social decision-making.

Key words social framing effect, resting-state fMRI, functional connectivity, right temporoparietal junction, moral network